E. Slobodyansky

#21

1600

# ENTERED

```
 3 <110> APPLICANT: Lanes, Olav
 4        Willasen, Nils Peder
 5        Guddal, Per Henrik
 6        Gjellesvik, Dag Rune
 8 <120> TITLE OF INVENTION: COD URACIL-DNA GLYCOSYLASE, GENE CODING THEREFORE,
RECOMBINANT DNA
 9        CONTAINING SAID GENE OR OPERATIVE PARTS THEREOF, A METHOD FOR PREPARING SAID
10        PROTEIN AND THE USE OF SAID PROTEIN OR SAID OPERATIVE PARTS THEREOF IN
11        MONITORING OR CONTROLLING PCR
13 <130> FILE REFERENCE: U 013209-3
15 <140> CURRENT APPLICATION NUMBER: 09/758,017B
16 <141> CURRENT FILING DATE: 2001-01-10
18 <150> PRIOR APPLICATION NUMBER: 2000 5428
19 <151> PRIOR FILING DATE: 2000-10-27
21 <150> PRIOR APPLICATION NUMBER: 2000 0163
22 <151> PRIOR FILING DATE: 2000-01-12
24 <160> NUMBER OF SEQ ID NOS: 22
26 <170> SOFTWARE: PatentIn version 3.1
28 <210> SEQ ID NO: 1
29 <211> LENGTH: 1283
30 <212> TYPE: DNA
31 <213> ORGANISM: GADUS MORHUA
33 <220> FEATURE:
34 <221> NAME/KEY: CDS
35 <222> LOCATION: (18)..(920)
36 <223> OTHER INFORMATION:
W--> 38 <400> 1
39 gacatccgct tgcaaat atg ttg ttc aag tta ggg tta tgc caa aga tgc      50
40                 Met Leu Phe Lys Leu Gly Leu Cys Gln Arg Cys
41                  1           5               10
43 ata tca tca aat cgg gty tta cca ggt tta cta att ccc caa act tta      98
W--> 44 Ile Ser Ser Asn Arg Xaa Leu Pro Gly Leu Leu Ile Pro Gln Thr Leu
45          15              20              25
47 tgt ttt tct aaa tta atg aag ata acg ccg aag aaa ctg agg tcc tca     146
48 Cys Phe Ser Lys Leu Met Lys Ile Thr Pro Lys Lys Leu Arg Ser Ser
49          30              35              40
51 aat gtg gaa caa aag acg tca tcg cca cag ctt tca gtg gag cag ctg     194
52 Asn Val Glu Gln Lys Thr Ser Ser Pro Gln Leu Ser Val Glu Gln Leu
53          45              50              55
55 gaa aga atg gcc aaa aat aag aaa gca gcg ctt gac aag att aga gca     242
56 Glu Arg Met Ala Lys Asn Lys Lys Ala Ala Leu Asp Lys Ile Arg Ala
57 60              65              70              75
59 aaa gca acg cct gca ggt ttc gga gag act tgg aga aga gag ctg gct     290
60 Lys Ala Thr Pro Ala Gly Phe Gly Glu Thr Trp Arg Arg Glu Leu Ala
```

```
61              80              85              90
63 gca gag ttt gaa aag cca tac ttc aaa caa ttg atg tcc ttt gta gct    338
64 Ala Glu Phe Glu Lys Pro Tyr Phe Lys Gln Leu Met Ser Phe Val Ala
65              95              100             105
67 gat gag agg agc cgt cac acc gtc tac cca ccg gct gat caa gtg tac    386
68 Asp Glu Arg Ser Arg His Thr Val Tyr Pro Pro Ala Asp Gln Val Tyr
69         110             115             120
71 agt tgg aca gag atg tgt gac att caa gat gtg aaa gta gtg att cta    434
72 Ser Trp Thr Glu Met Cys Asp Ile Gln Asp Val Lys Val Val Ile Leu
73     125             130             135
75 ggc cag gac cct tac cac ggt ccc aac caa gca cat gga ctc tgt ttc    482
76 Gly Gln Asp Pro Tyr His Gly Pro Asn Gln Ala His Gly Leu Cys Phe
77 140             145             150             155
79 agt gtg caa aag cca gtt ccc cct ccc ccc agt ctc gtg aac ata tac    530
80 Ser Val Gln Lys Pro Val Pro Pro Pro Ser Leu Val Asn Ile Tyr
81              160             165             170
83 aaa gaa ttg tgt acc gac att gat ggc ttc aag cat cct gga cat gga    578
84 Lys Glu Leu Cys Thr Asp Ile Asp Gly Phe Lys His Pro Gly His Gly
85         175             180             185
87 gat cta agc gga tgg gca aaa caa ggg gtg ctg ctg ctt aac gcg gtg    626
88 Asp Leu Ser Gly Trp Ala Lys Gln Gly Val Leu Leu Leu Asn Ala Val
89         190             195             200
91 ctg acc gtg cgg gcc cat cag gcc aac tcc cac aag gac aga ggc tgg    674
92 Leu Thr Val Arg Ala His Gln Ala Asn Ser His Lys Asp Arg Gly Trp
93     205             210             215
95 gag acc ttc acc gac gct gtg atc aag tgg ctg agc gtc aac cgg gaa    722
96 Glu Thr Phe Thr Asp Ala Val Ile Lys Trp Leu Ser Val Asn Arg Glu
97 220             225             230             235
99 gga gtg gtt ttc ctg ttg tgg ggc tca tac gcc cat aag aag gga gcg    770
100 Gly Val Val Phe Leu Leu Trp Gly Ser Tyr Ala His Lys Lys Gly Ala
101             240             245             250
103 acc atc gac agg aaa cgt cac cat gtc ttg caa gct gtt cat cca tct   818
104 Thr Ile Asp Arg Lys Arg His His Val Leu Gln Ala Val His Pro Ser
105         255             260             265
107 cct ttg tct gct cat cgt ggg ttc ctt ggt tgt aag cac ttc tcc aag   866
108 Pro Leu Ser Ala His Arg Gly Phe Leu Gly Cys Lys His Phe Ser Lys
109     270             275             280
111 gct aac ggg ctg ctg aaa cta tct ggg acg gag cct ata aac tgg aga   914
112 Ala Asn Gly Leu Leu Lys Leu Ser Gly Thr Glu Pro Ile Asn Trp Arg
113     285             290             295
115 gca ctc taactcttta tgctgcctta tactgttaac tgttttaaga tgaacatcac   970
116 Ala Leu
117 300
119 actatatttt ctacagcttt tccaagttca aaccaatcta taagctttca tttgtctttt  1030
121 ggaatgatgc tgcttttggt cggttttaga tacttaaaac actttaccac tctgccatgt  1090
123 tgactcatgt tcagtcaata taactttcac aacttgaaca aaaatgttat tttataattg  1150
125 attatattct gtacattaaa gattgttttt ttcccaggct gtttcatagg tactaggata  1210
127 ttaaactgtt attaacctat tttccatgat gtcaactgct taagttttta tgcagaaata  1270
129 aattatatat tta                                                    1283
```

```
132 <210> SEQ ID NO: 2
133 <211> LENGTH: 301
134 <212> TYPE: PRT
135 <213> ORGANISM: GADUS MORHUA
137 <220> FEATURE:
138 <221> NAME/KEY: misc_feature
139 <222> LOCATION: (17)..(17)
140 <223> OTHER INFORMATION: The 'Xaa' at location 17 stands for Val.
142 <400> SEQUENCE: 2
144 Met Leu Phe Lys Leu Gly Leu Cys Gln Arg Cys Ile Ser Ser Asn Arg
145 1               5                   10                  15
```

W--> `148 Xaa Leu Pro Gly Leu Leu Ile Pro Gln Thr Leu Cys Phe Ser Lys Leu`

```
149             20                  25                  30
152 Met Lys Ile Thr Pro Lys Lys Leu Arg Ser Ser Asn Val Glu Gln Lys
153         35                  40                  45
156 Thr Ser Ser Pro Gln Leu Ser Val Glu Gln Leu Glu Arg Met Ala Lys
157     50                  55                  60
160 Asn Lys Lys Ala Ala Leu Asp Lys Ile Arg Ala Lys Ala Thr Pro Ala
161 65                  70                  75                  80
164 Gly Phe Gly Glu Thr Trp Arg Arg Glu Leu Ala Ala Glu Phe Glu Lys
165             85                  90                  95
168 Pro Tyr Phe Lys Gln Leu Met Ser Phe Val Ala Asp Glu Arg Ser Arg
169             100                 105                 110
172 His Thr Val Tyr Pro Pro Ala Asp Gln Val Tyr Ser Trp Thr Glu Met
173         115                 120                 125
176 Cys Asp Ile Gln Asp Val Lys Val Val Ile Leu Gly Gln Asp Pro Tyr
177     130                 135                 140
180 His Gly Pro Asn Gln Ala His Gly Leu Cys Phe Ser Val Gln Lys Pro
181 145                 150                 155                 160
184 Val Pro Pro Pro Ser Leu Val Asn Ile Tyr Lys Glu Leu Cys Thr
185             165                 170                 175
188 Asp Ile Asp Gly Phe Lys His Pro Gly His Gly Asp Leu Ser Gly Trp
189             180                 185                 190
192 Ala Lys Gln Gly Val Leu Leu Leu Asn Ala Val Leu Thr Val Arg Ala
193         195                 200                 205
196 His Gln Ala Asn Ser His Lys Asp Arg Gly Trp Glu Thr Phe Thr Asp
197     210                 215                 220
200 Ala Val Ile Lys Trp Leu Ser Val Asn Arg Glu Gly Val Val Phe Leu
201 225                 230                 235                 240
204 Leu Trp Gly Ser Tyr Ala His Lys Lys Gly Ala Thr Ile Asp Arg Lys
205             245                 250                 255
208 Arg His His Val Leu Gln Ala Val His Pro Ser Pro Leu Ser Ala His
209             260                 265                 270
212 Arg Gly Phe Leu Gly Cys Lys His Phe Ser Lys Ala Asn Gly Leu Leu
213         275                 280                 285
216 Lys Leu Ser Gly Thr Glu Pro Ile Asn Trp Arg Ala Leu
217     290                 295                 300
220 <210> SEQ ID NO: 3
221 <211> LENGTH: 1355
```

```
222 <212> TYPE: DNA
223 <213> ORGANISM: GADUS MORHUA
225 <220> FEATURE:
226 <221> NAME/KEY: CDS
227 <222> LOCATION: (90)..(992)
228 <223> OTHER INFORMATION:
W--> 230 <400> 3
231 gatggtttag gaggatagta ctttgacact ggttagcgaa ggggaaaacg gagttattgt      60
233 gcatatcgtt ttagccctac gtttaaaaa atg att ggt caa cag cat atc aac     113
234                                 Met Ile Gly Gln Gln His Ile Asn
235                                 1               5
237 tct ttc ttc tca cca gtt tca aaa aag aga gtt tca aag gaa tta ggt     161
238 Ser Phe Phe Ser Pro Val Ser Lys Lys Arg Val Ser Lys Glu Leu Gly
239     10              15                  20
241 aaa acc gaa aag cat gcc gaa gaa gtt cag ata acg ccg aag aaa ctg     209
242 Lys Thr Glu Lys His Ala Glu Glu Val Gln Ile Thr Pro Lys Lys Leu
243 25              30                  35                  40
245 agg tcc tca aat gtg gaa caa aag acg tca tcg cca cag ctt tca gtg     257
246 Arg Ser Ser Asn Val Glu Gln Lys Thr Ser Ser Pro Gln Leu Ser Val
247             45                  50                  55
249 gag cag ctg gaa aga atg gcc aaa aat aag aaa gca gcg ctt gac aag     305
250 Glu Gln Leu Glu Arg Met Ala Lys Asn Lys Lys Ala Ala Leu Asp Lys
251             60                  65                  70
253 att aga gca aaa gca acg cct gca ggt ttc gga gag act tgg aga aga     353
254 Ile Arg Ala Lys Ala Thr Pro Ala Gly Phe Gly Glu Thr Trp Arg Arg
255         75                  80                  85
257 gag ctg gct gca gag ttt gaa aag cca tac ttc aaa caa ttg atg tcc     401
258 Glu Leu Ala Ala Glu Phe Glu Lys Pro Tyr Phe Lys Gln Leu Met Ser
259     90                  95                  100
261 ttt gta gct gat gag agg agc cgt cac acc gtc tac cca ccg gct gat     449
262 Phe Val Ala Asp Glu Arg Ser Arg His Thr Val Tyr Pro Pro Ala Asp
263 105                 110                 115                 120
265 caa gtg tac agt tgg aca gag atg tgt gac att caa gat gtg aaa gta     497
266 Gln Val Tyr Ser Trp Thr Glu Met Cys Asp Ile Gln Asp Val Lys Val
267                 125                 130                 135
269 gtg att cta ggc cag gac cct tac cac ggt ccc aac caa gca cat gga     545
270 Val Ile Leu Gly Gln Asp Pro Tyr His Gly Pro Asn Gln Ala His Gly
271                 140                 145                 150
273 ctc tgt ttc agt gtg caa aag cca gtt ccc cct ccc ccc agt ctc gtg     593
274 Leu Cys Phe Ser Val Gln Lys Pro Val Pro Pro Pro Ser Leu Val
275             155                 160                 165
277 aac ata tac aaa gaa ttg tgt acc gac att gat ggc ttc aag cat cct     641
278 Asn Ile Tyr Lys Glu Leu Cys Thr Asp Ile Asp Gly Phe Lys His Pro
279     170                 175                 180
281 gga cat gga gat cta agc gga tgg gca aaa caa ggg gtg ctg ctg ctt     689
282 Gly His Gly Asp Leu Ser Gly Trp Ala Lys Gln Gly Val Leu Leu Leu
283 185                 190                 195                 200
285 aac gcg gtg ctg acc gtg cgg gcc cat cag gcc aac tcc cac aag gac     737
286 Asn Ala Val Leu Thr Val Arg Ala His Gln Ala Asn Ser His Lys Asp
```

RAW SEQUENCE LISTING
PATENT APPLICATION:  US/09/758,017B

DATE: 05/30/2003
TIME: 13:47:41

Input Set : A:\EP.txt
Output Set: N:\CRF4\05302003\I758017B.raw

```
287                  205              210              215
289 aga ggc tgg gag acc ttc acc gac gct gtg atc aag tgg ctg agc gtc       785
290 Arg Gly Trp Glu Thr Phe Thr Asp Ala Val Ile Lys Trp Leu Ser Val
291              220              225              230
293 aac cgg gaa gga gtg gtt ttc ctg ttg tgg ggc tca tac gcc cat aag       833
294 Asn Arg Glu Gly Val Val Phe Leu Leu Trp Gly Ser Tyr Ala His Lys
295          235              240              245
297 aag gga gcg acc atc gac agg aaa cgt cac cat gtc ttg caa gct gtt       881
298 Lys Gly Ala Thr Ile Asp Arg Lys Arg His His Val Leu Gln Ala Val
299      250              255              260
301 cat cca tct cct ttg tct gct cat cgt ggg ttc ctt ggt tgt aag cac       929
302 His Pro Ser Pro Leu Ser Ala His Arg Gly Phe Leu Gly Cys Lys His
303 265              270              275              280
305 ttc tcc aag gct aac ggg ctg ctg aaa cta tct ggg acg gag cct ata       977
306 Phe Ser Lys Ala Asn Gly Leu Leu Lys Leu Ser Gly Thr Glu Pro Ile
307              285              290              295
309 aac tgg aga gca ctc taactcttta tgctgcctta tactgttaac tgttttaaga      1032
310 Asn Trp Arg Ala Leu
311              300
313 tgaacatcac actatatttt ctacagcttt tccaagttca aaccaatcta taagctttca     1092
315 tttgtctttt ggaatgatgc tgcttttggt cggttttaga tacttaaaac actttaccac     1152
317 tctgccatgt tgactcatgt tcagtcaata taactttcac aacttgaaca aaaatgttat     1212
319 tttataattg attatattct gtacattaaa gattgttttt ttcccaggct gtttcatagg     1272
321 tactaggata ttaaactgtt attaacctat tttccatgat gtcaactgct taagtttta      1332
323 tgcagaaata aattatatat tta                                            1355
```

326 <210> SEQ ID NO: 4
327 <211> LENGTH: 301
328 <212> TYPE: PRT
329 <213> ORGANISM: GADUS MORHUA
331 <400> SEQUENCE: 4

```
333 Met Ile Gly Gln Gln His Ile Asn Ser Phe Phe Ser Pro Val Ser Lys
334 1               5               10              15
337 Lys Arg Val Ser Lys Glu Leu Gly Lys Thr Glu Lys His Ala Glu Glu
338              20              25              30
341 Val Gln Ile Thr Pro Lys Lys Leu Arg Ser Ser Asn Val Glu Gln Lys
342          35              40              45
345 Thr Ser Ser Pro Gln Leu Ser Val Glu Gln Leu Glu Arg Met Ala Lys
346      50              55              60
349 Asn Lys Lys Ala Ala Leu Asp Lys Ile Arg Ala Lys Ala Thr Pro Ala
350 65              70              75              80
353 Gly Phe Gly Glu Thr Trp Arg Arg Glu Leu Ala Ala Glu Phe Glu Lys
354              85              90              95
357 Pro Tyr Phe Lys Gln Leu Met Ser Phe Val Ala Asp Glu Arg Ser Arg
358              100             105             110
361 His Thr Val Tyr Pro Pro Ala Asp Gln Val Tyr Ser Trp Thr Glu Met
362          115             120             125
365 Cys Asp Ile Gln Asp Val Lys Val Val Ile Leu Gly Gln Asp Pro Tyr
366      130             135             140
369 His Gly Pro Asn Gln Ala His Gly Leu Cys Phe Ser Val Gln Lys Pro
```

RAW SEQUENCE LISTING ERROR SUMMARY     DATE: 05/30/2003
PATENT APPLICATION: **US/09/758,017B**     TIME: 13:47:42

Input Set : **A:\EP.txt**
Output Set: **N:\CRF4\05302003\I758017B.raw**

**Please Note:**
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the
Sequence Listing to ensure that a corresponding explanation is presented in the <220>
to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; Xaa Pos. 17
Seq#:2;. Xaa Pos. 17

**Invalid <213> Response:**
Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:14,21,22

**VERIFICATION SUMMARY**　　　　　　　DATE: 05/30/2003
PATENT APPLICATION:　**US/09/758,017B**　　TIME: 13:47:42

Input Set : **A:\EP.txt**
Output Set: **N:\CRF4\05302003\I758017B.raw**

```
L:38 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:1,Line#:36
L:44 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1 after pos.:98
L:148 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2 after pos.:16
L:230 M:258 W: Mandatory Feature missing, <223> Blank for SEQ#:3,Line#:228
```